Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at http://about.jstor.org/participate-jstor/individuals/early-journal-content.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

A SCALE FOR MEASURING THE GENERAL MERIT OF ENGLISH COMPOSITION IN THE SIXTH GRADE

FREDERICK S. BREED AND F. W. FROSTIC
Department of Education, University of Michigan

INTRODUCTION

The movement in scientific educational procedure has made rapid progress in the past five or six years. One phase of this movement that has particularly attracted wide attention is the measurement of abilities of pupils in elementary-school subjects. The demand for definite objective measurements of educational products is becoming more insistent in all departments of school work. Administrators, supervisors, teachers, and students of education generally are no longer satisfied with mere opinion regarding educational progress and efficiency. They are seeking for information of a definite character concerning the work of their schools.

The interest manifested in such measurements is prompted by a desire to determine where we are, what progress we are making, and where we ought to be. And the "ought" in this case is related to the "is." Standardization, which has for its aim the determination of the proper requirements in the various school subjects, should rest, wherever possible, on measurement. If an administrator finds that a particular class spells correctly 88 per cent of the words in a list selected from Column O of the Ayres spelling scale, it is clear that he has a definite objective measurement of the spelling achievement of that class. If he finds that a class advances from quality 11 to quality 12 on the Thorndike handwriting scale, it is clear that he has a definite measure of the improvement of this class. Finally, it seems clear that by means of just such measurements as these the most satisfactory standards of educational attainments are to be established. What a pupil ought to do should be determined on the basis of what he can do. Standards must be dependent on actual measurement and not on arbitrary judgment.

Among the various school subjects, English composition is generally conceded to be one of the most difficult to measure. The best-known means that have been prepared for the measurement of this subject are the Hillegas-Thorndike scale and the Harvard-Newton scales. The Hillegas-Thorndike scale measures the "quality of English composition by young people," and by quality is meant general merit. The Harvard-Newton scales, four in number, are designed to measure the ability of upper grammar-grade pupils in each of the four types of composition.

The scale described and reproduced below has been devised to measure the general merit of English composition of sixth-grade children. It is not a scale in the same sense as the instrument which the grocer uses for weighing sugar is a scale. It is rather like the supposititious scale described by Thorndike[1] in his monograph, "The Measurement of Achievement in Drawing." "If we had a series of blocks of lead," says Thorndike, "weighing 2.4, 3.9, 5.7, 6.5, 7.8, 8.6, 10.5, 11.8, 12.6, 13.5, 14.4, 16, and 17 pounds respectively, they would be a scale or graded series of standards for weight substantially as these drawings are a graded series of standards of children's drawings."

This kind of scale is not original with scientific students of education. It is well known in other scientific fields. Some readers may be familiar with the scale of hardness used in mineralogy, or with the hemoglobin scale used by physicians, or with the tintometer of the industrial arts. The scale of hardness is a series of mineral specimens ranging through ten steps of hardness. The hardness of each of the specimens of the scale has been carefully determined. The hardness of any unknown mineral may be tested by comparing it with the scale. When a mineral on the scale is found that resists abrasion to about the same degree as the unknown, the hardness of that member of the scale is assigned to the mineral tested. The hemoglobin scale of the physician is used in a similar manner. A drop of blood is allowed to fall on standard blotting-paper, and the stain is matched against a series of tints on a lithographic scale.

[1] *Teachers College Record*, XIV, 352.

A Scale For Measuring the General Merit of English Composition in the Sixth Grade

Frederick S. Breed and F. W. Frostic
University of Michigan

The numbers above the samples represent amounts of merit.

.2

The hanjict shop for there
there was so many in it. After
a little they it going. And
they to pleace were the the picine
was. They all get out of the
hounce and but there thing
down
on the table and rain out to
play. After they were platy
the had there lunce. And
They had a very nice time.

2.7

When the engine r stop, one
of the boy took his shoes and
off
stcocking˅and got out into the
river and bushe the bout a little
so as to stare the enginer
a going, when that had
stared the bowt went aright.
And they went rideing
around the river where
having a nice time went
one of the girl saw a water—
lillies and they try to pick when she
fell in the river but she got
aright, her cloth were wet
some but they soon try, and
she got her water lillies.
When they were throu-
ght rideing they got out and
went to the bank and
had their lounce, After
that they play games,
which then they began
to pick flowers in the woold.
But when they went home
they had many beautiful flower
to take their mothers. But
when the departed they
all to their teacher we all
have like you this year and
hope you will teach again
next year. And we all are glady
we have pass this year.

3.7

They were near the park when the ~~enjing~~
engin stoped. They pattled ashore and went on
there
to the parck. ~~There~~ t There, ~~the~~ was a man who
had another f boat. ~~At~~ After a while they
asked the man if he would tow the boat lack
when he went and he said that he would . When
evening came the man told the teacher that he
was going f back. The teacher ~~glat~~ gathered the
children together and went to there boat where
on
the man had already tied the top rope ~~one~~ the
were towed
boat. They all got in and ~~wont~~ back ~~t~~ up the
river. Then they departed and went home happily.

4.7

Jack the one who was runing
the launch said, " lets take these pales
and push the launch to shore. Yes! Yes!
let do said all; we took the poles
and tu) pushed it to shore. Jack
got out and looked around. He saw
a house over yonder, he went over
to see if they had any gasoline,
when. he got there and ask they
had none. When he came back
we were suprised to (she) see
him without the gasoline. He
said, "lets go out here and walk
the rest of the way. We (w) all
said yes, because it was only about
one and a half (ba) blocks or so.
When we arrived at the wood
we all ran to get a the
tables set and get every thing
ready, that so went we got
through picking a few we coald
eat our lunch. When we got
enough (w) violets, of purple, white
and yellow, dutcchmans breechs,
an liles and etc. we ate our
luanch and had a fine time.
We got ready to go home
and a big cloud of black was
over us, we got ready just the
same and had to push it alone the way hon:

5.7

After a time floating on the river.They got
it started and away they went, It was two
oclock when they got to the picnic grounds.
All played games till 3oclock, then they
had lunch, and played awhile longer, then
got ready to come back. John took the
baskets to the river to puttin the boat
but it was gone. one of the boys went
up the river a little ways, but
stayed to long, The children thought some
one stole it. They were all there, but where was
Jack? Henery heard the chucking of a boat,
He ran to the river, and saw Jack coming
around the curve, He gave Jack a scolding
for taking the boat. So they loaded the things in
and started for home. They all landed
safe, and were glad that the engine did
not buck, that ended the picnic.

7.0

"Oh," crief Jhoney, "what shall we
do. The teacher who had been
shouting with the children, turned
pale for just ahead of them was
a swift currant which, she knew,
would (lu) hurrel the launch
up side down and drownd them all.
Some of the little girls began to
cry but the teacher relized that
something must be done at once.
She was a brave and gritty little lady
and so she resolved to do
a dangerous thing. She knew that
if the boat could be turned on some
other course that they might
see a boat and be saved but
no body ever ventured in this
dangerous place.
~~There was a rope in~~ A little
ways from the launch was a
~~be~~ long board. (~~If they could~~)
floating around.
If they could but reach this it
might mean safty for they could
push the boat out of this course.
There was a rope in the launch
and the teacher tied it around her
waist and had four of her strongest
boys to hold on the other end.
The teacher jumped over board for
she could swim quite good.
Three times the board slipped
out of her hands and she was almost
discouraged. when at last she got it.
A shout of joy arose from the boat
~~and~~ but the teacher had no
time to join them for in two
more minutes the boat would be
in the currant.
She got on board and ~~in~~
before they could realize what
was happening she was pusshing
the boat out of its course and some
of the big boys were helping her.
At last the boat moved into other

8.0

They had run out of gasoline. Where
they were it was ten feet deep and no
one could swim.
What could they do with all these child
ren, crying and shreiking with the fear of
not being able to have the picnic.
The man that ran the engine told the
others that they must wait for someone to
come along and take them to the dock. Then, the
children were getting nervous about
staying over night.
Suddenly a bright thought came to
the teacher. She said, "Children, can't we have
a picnic on the boat? There is plenty of sit-
ting down games, I am shure. "This Idea pleas
ed the children and soon they were very busy
laughing. Soon as the big steamer 'Armada'
went past and offered to take them in but
the children were having so good a time
that they asked the 'Armada' to come back at
five o'clock.
They had their picnic lunch on the
boat at four o'clock, and by five they were
willing to go back.
They afterwards said it was more
fun than all the other picnics put to-

9.0

"What! the engine stopped, well I
declare," ex claimed the teacher, "and we
time
were planning such a fine˅ too," she finished
I am very sorry this has happed
madam," said the boatman," but we are
in no great danger, as I have a friend
near here who will help me get ashore."
At this the boatman took a
shining white whistle from his pocket
and blew three long shrill blasts,
which I suppose was some signal for a
heavily built man was soon on the
bank and after finding what was wanted
he was soon in his boat house after his row boat.
Soon every one was safely landed
on the opposite shore. and to judge by
their faces enjoying themselves very
much.
The boat was fixed in a short
time and they all got in and had a
happy time riding.
They all thought this landing place
was fully as nice as the other and
it fact as good as any other.

9.7

"What is the matter"? asked Miss Green.
the teacher of the school, to the man who
was running the little launch. "I'm
sorry mam," he said, "But the gasoline
has given out." Some of the more timid
children began to cry. "Hush," said the
teacher a little sternly, "We cannot let
this accident spoil our picnic." "Why
can't we have our lunch right in the
boat?" cried Mary, Brightly. "That is
a fine idea, Mary," the teacher said
gratefully. With the help of the pupils
Miss Green managed to set the lunch
on the floor of the boat. After lunch
their spirits began to rise, and they amused
themselves by telling stories.
All this while they were drifting
along quietly. "There is a launch
coming towards us!" exclaimed a small
boy. Miss Green looked and sure
enough there was a launch heading
directly towards them. They were towed
back home by this boat. That was
a fine picnic after all, "chorused the
children as they stepped on firm
land once more.

The scale indicates hemoglobin content from 10 to 100 per cent. The tintometer is used in like fashion to determine the tints of solutions and fabrics.

In the same manner we have in this composition scale a series of samples whose values have been carefully determined by the average judgment of over sixty individuals, nearly all of whom were experienced teachers, twenty of them specialists in English, and the remainder students of education. These samples have been arranged in ascending order, 0.2, 2.7, 3.7, 4.7, 5.7, 7.0, 8.0, 9.0, and 9.7 being the values assigned. Just as the student of mineralogy can determine the hardness of a mineral more accurately with the hardness scale, and the physician can estimate the amount of color in the blood with greater precision by the use of the hemoglobin scale, so the teacher can determine the merit of a sample of English composition with greater precision by comparing it with a graded series of evaluated compositions such as the present scale.

CONSTRUCTION OF THE SCALE

The construction of this scale grew out of an attempt to survey the instruction in English in a selected sixth grade in each of ten Michigan cities. The following letter was sent to the superintendents of ten school systems inviting their co-operation in the work:

DEAR SIR:

With the assistance of advanced students in the Department of Education of the University of Michigan, I am planning to make a limited survey of the work of the sixth grade in each of ten Michigan public-school systems. It is our desire that the school systems co-operating in the work be of varying size, and, where there are a number of sixth grades, that the grade selected for testing be, as nearly as can be judged, as good as any in the city. It is our intention to submit to the principal and teacher, or teachers, with the consent of the superintendent, not over six tests, with directions for giving. The tests are limited to six because we wish to confine ourselves to those subjects in which standard scores or scales of measurement have been developed. The subjects are: reading, arithmetic, writing, English composition, spelling, and drawing. After the work has been done by the pupils, it will be forwarded to the University and scored by means of the scales now available in the subjects named. The purpose of the investigation is threefold: (1) to demonstrate the use of scales of measurement, (2) to secure for the superintendent a comparative estimate of the work of a sample grade in his schools, (3) to make a study of the scales of measurement and their application.

It is our desire to go ahead with the work as soon as the schools approve and the tests can be prepared. If you feel that the survey as outlined would be beneficial to you and your schools, we should be pleased to have you join with us. An early reply will be appreciated.

The names of the towns and cities in which these systems of schools were located, with populations according to the *Michigan Teachers' Directory* for 1915, are as follows:

South Haven	4,500	Port Huron	22,500
Three Rivers	5,500	Saginaw, E.S.	30,000
Coldwater	6,000	Lansing	42,000
Owosso	10,000	Kalamazoo	48,000
Ann Arbor	16,000	Grand Rapids	112,571

Favorable replies were received from all of the ten. The following composition test, with accompaning directions for giving, was then sent to the superintendents along with the tests in the five other branches included in the survey.

COMPOSITION TEST

When the sixth grade is distributed through two or more rooms see that this test is given in all these rooms during the same session.

Uniform paper (which has been supplied) and pens should be distributed and ready for use before the teacher begins the test. See that *each sheet* is headed with (*a*) the pupil's name, (*b*) the grade, (*c*) the school, (*d*) the city, (*e*) the date.

Say to the children: "I am going to read you the first part of a story called 'The Picnic,' and when I am through I want you to write the rest of it for me. Listen carefully."

THE PICNIC

It was a beautiful day in June, and the pupils were busily planning for a picnic which was to take place that afternoon in the park a mile down the river.

"Now, John, don't forget to bring your spoon and tin cup, and, Mary, be sure that you have enough sandwiches," said the teacher. After a few words to some of the other pupils who were going to the picnic, she dismissed them, all eager for the coming event, and all promising to be on hand at one o'clock, at the landing, ready for a jolly time.

One o'clock found them all assembled, with baskets packed and happy faces. The boat in which they were to start was a sturdy little gasoline launch, with plenty of room in it for all. They stowed themselves aboard, and shouted to be off. Soon they were chugging merrily down the stream.

Everything seemed to be going along smoothly, when suddenly the engine stopped running!

"Now, I want you to use your imagination and make up *the rest* of the story for yourself. Think of what *you* would have done if you had been on the boat when it stopped, and what other adventures the pupils might have had during the rest of the afternoon. I shall give you *twenty minutes* in all to think of your story and to write it, but do not start writing until you have thought of something that will make a good story. Are there any questions?"

Answer any questions which may arise, but be careful not to give the pupils any hints which might suggest to them a possible ending for the story. Be sure that the pupils understand that a *reproduction* of the story read is *not* wanted. Then, say, "You may begin."

Keep the time accurately from this point, and at the end of exactly *twenty minutes*, say, "Stop writing," and collect the papers.

Please indicate the sex of the pupil when the same is not clear from the name.

By this means 481 samples of English composition were obtained on the same subject and under relatively similar conditions. Since the superintendents were asked to submit the test to a sixth grade "as good as any in the city," it is fair to suppose that many of them submitted it to the better classes, if not to the best, and that the samples therefore represent a range of ability somewhat above the normal for this grade. After the samples had been numbered and were ready to be scored, it seemed possible to build out of this material a scale that would score the results from these several schools more satisfactorily than the Hillegas-Thorndike scale.

The 481 samples were thereupon submitted to three competent judges, who graded them independently into ten groups according to their general merit as samples of English composition. The samples were arranged in such a manner that the poorest specimens were placed in group 1, next to the poorest in group 2, and so on by approximately equal steps to the best group, which was marked 10. The reports of the three judges were then tabulated and a set of 70 samples chosen from the 481 for further investigation. These samples were selected according to agreement in the reports of the judges, i.e., those samples were selected for further use which showed an agreement in the reports of two or more judges in placing them in a given group. As might be expected, there were very few extremely poor samples and very few of the highest quality. Almost all of the samples at the extremes of the distribution were included in the list of 70.

This set of 70 samples therefore represented the total range of composition ability of the grade measured. These samples were typewritten in such a manner as to reproduce as exactly as possible the form of the compositions, handwriting alone being excluded. This meant that margins, paragraphing, relative length of lines, and erasures, as well as errors in spelling, capitalization, and punctuation, were reproduced. The 70 typewritten samples were then submitted to 20 judges, students of education, most of whom were experienced teachers, with the following instructions:

Directions to Judges

1. Read the accompanying "Composition Test," showing the conditions under which the compositions were produced.

2. Read the samples carefully, arranging them in five groups according to their general merit as samples of English composition. Put those which appear to be the best in group (A), the next in rank in group (B), and so on in groups (C), (D), and (E). Do not try to put the same number in each group, but simply group those which seem nearest in rank for convenience in step 3.

3. Then, beginning with the poorest group, find the poorest sample, then by comparison the next to the poorest and so on until the best of all the samples in all the groups is found. It is best to compare but two samples at once and decide their relative positions by judging whether one is better or worse than the other. You should now have the seventy samples arranged in the order of merit with the poorest on the bottom and the best on the top.

4. Compare the samples once more and make any corrections you think necessary in the relative rank of each.

5. Fill out the following table from your arrangement of samples, marking the *poorest sample* 1, and so on up to the best, which should be marked 70.

Number of sample	Rank assigned
1	....
2	....
3	....
4	....
5	....
etc., to	
70	....

6. Please do not confer with anyone on any part of the work.
7. Kindly return the compositions with your report in the inclosed stamped and addressed envelope.
8. Write your name and address below and estimate the time required to do the work.

Name....................Address.................. Time..........

From the reports of these judges Table I was prepared, showing the positions assigned the various samples. Only 25 of the 70 samples appear in the table, these being the 25 that were subsequently used in further investigation. The 25 samples were selected by a method similar to that used in the selection of the final samples of the scale which is described more fully on a later page. From the set of 70 samples it was found that at least 14 separate scales could be selected in which a scale step varied from 0.75 to 1.25 P.E. Of the 70 samples, 35 were included in one or more of these scales. Of these 35, 10 were rejected. These 10 samples were located at intermediate steps where there were other samples with a greater frequency of appearance in the 14 scales. Thus there were left for further study 25 samples.

From Table I, Table II was compiled showing the number of better judgments. The number of better judgments in each case was determined by comparing the positions assigned each sample with the positions assigned every other sample in the set. For example, if we compare sample 421 with sample 275 in Table I, it will be seen that sample 421 was judged better than sample 275 six times, that is, 6 of the 20 judges thought sample 421 better than sample 275.

In order to secure a more accurate rating through greater competency of judges, these 25 samples were submitted to 20 special teachers of English, most of whom were teachers in Michigan high schools, with the same directions for judging the samples and recording the results as were sent to the first set of 20 judges, except that, since there were but 25 samples in this set, they were instructed, in arranging the samples in order, to call the poorest sample 1 and the best 25. The results from these judges appear in Tables III and IV, Table IV being prepared from Table III in the same manner as Table II was prepared from Table I.

TABLE I

Positions Assigned by 20 Judges to 25 of the Set of 70 Samples Considered by These Judges

Judges	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
312	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1
289	2	3	2	3	4	2	2	4	6	2	3	2	5	5	4	2	4	4	2	6
303	11	2	7	10	11	29	13	3	10	4	2	5	12	7	2	4	10	3	9	8
195	8	10	20	14	10	8	27	7	7	9	10	15	21	14	13	5	18	6	11	7
421	10	20	21	5	17	6	5	37	14	5	19	11	11	13	17	7	12	18	25	24
275	23	23	28	13	25	18	11	11	12	10	22	14	18	18	16	17	27	14	6	9
168	25	15	36	18	13	27	19	20	11	18	14	16	20	20	20	9	14	16	7	34
360	36	12	18	39	27	45	22	27	24	16	13	27	34	29	21	10	19	23	37	13
324	35	13	12	22	16	41	37	12	16	15	12	40	55	30	15	49	50	15	28	20
73	26	28	13	16	32	16	20	30	21	20	27	33	30	19	47	65	32	55	48	28
454	38	45	24	46	23	44	14	18	25	11	45	34	33	24	22	24	51	22	20	32
306	34	22	30	35	28	42	42	25	17	29	24	20	31	35	25	43	33	21	34	12
166	41	34	42	31	31	24	25	22	29	25	36	53	40	49	28	23	20	26	44	65
30	27	38	16	26	41	53	46	23	39	34	40	23	19	47	24	67	28	50	62	37
228	39	62	45	37	54	26	47	55	56	50	62	42	53	31	54	28	57	27	24	57
63	46	50	51	43	59	39	57	36	51	54	51	43	24	22	66	68	64	54	40	54
89	58	66	43	21	56	50	48	52	37	56	64	51	52	56	44	61	59	52	50	44
287	63	67	57	62	62	62	51	48	42	60	66	55	35	61	38	44	68	60	36	58
385	70	65	70	70	64	70	65	59	59	46	68	68	43	68	39	41	42	70	53	50
126	64	39	67	50	47	54	50	54	64	70	41	58	59	60	43	63	67	37	65	62
257	69	37	64	65	49	52	62	65	57	65	37	38	51	65	64	57	66	64	47	61
342	67	54	40	63	65	66	63	57	65	55	53	67	63	69	62	33	37	61	69	53
212	66	68	65	64	57	65	58	62	70	41	67	47	67	54	63	70	70	44	56	56
448	65	63	68	66	67	57	61	58	63	53	63	59	69	67	65	69	43	67	68	69
8	68	70	60	68	70	59	68	68	67	64	70	69	65	70	70	66	54	69	67	67

TABLE II

NUMBER OF "BETTER" JUDGMENTS OF 20 JUDGES WITH A SET OF 70 SAMPLES

Only 25 of the 70 Samples Are Here Recorded. The Table Should Be Read, "None of the Judges Considered Sample 312 Better than Sample 289," etc.

Samples	289	303	195	421	275	168	360	324	73	454	306	166	30	228	63	89	287	385	126	257	342	212	448	8
312	0	0	0	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....
289	....	5	0	0	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....
303	....	....	5	5	3	1	0	0	1	0	0	1	0	1	0	....	....	....	....	....	....	....	....	....
195	....	....	....	8	4	4	2	1	2	1	0	1	1	0	0	....	....	....	....	....	....	....	....	....
421	....	....	....	....	6	7	5	8	2	2	2	1	2	0	0	0	....	....	....	....	....	....	....	....
275	....	....	....	....	....	5	5	5	2	2	0	1	1	0	0	0	....	....	....	....	....	....	....	....
168	....	....	....	....	....	....	5	8	5	4	2	0	2	1	0	0	0	....	....	....	....	....	....	....
360	....	....	....	....	....	....	....	11	8	9	7	4	5	3	2	0	1	0	....	....	....	....	....	....
324	....	....	....	....	....	....	....	....	8	6	5	5	4	4	4	2	2	3	0	2	2	0	1	0
73	....	....	....	....	....	....	....	....	....	5	11	7	5	3	2	2	3	2	3	2	1	0	....	....
454	....	....	....	....	....	....	....	....	....	....	11	5	8	2	4	0	0	1	0	2	1	0	....	....
306	....	....	....	....	....	....	....	....	....	....	....	7	5	4	0	1	0	1	0	0	1	0	0	....
166	....	....	....	....	....	....	....	....	....	....	....	....	7	5	5	3	3	1	1	1	2	1	0	....
30	....	....	....	....	....	....	....	....	....	....	....	....	....	4	3	4	2	2	1	5	1	2	0	1
228	....	....	....	....	....	....	....	....	....	....	....	....	....	....	9	5	4	3	5	5	4	2	1	1
63	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	10	4	5	7	6	5	4	3	2
89	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	3	5	5	7	6	4	4	1
287	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	5	9	6	7	5	5	1
385	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	12	13	11	11	11	5
126	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	11	5	6	3	3
257	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	10	9	4	4
342	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	9	8	2
212	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	8	6
448	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	5
8	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....

TABLE III

Positions Assigned to 25 Samples by 20 Judges

Judges	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
312	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
289	2	1	2	2	2	2	2	2	2	2	2	2	3	2	2	3	2	2	2	2
303	7	3	3	4	3	3	3	3	3	3	3	3	4	3	3	4	3	3	3	4
195	3	6	4	6	4	5	9	5	4	6	5	4	2	4	5	2	7	4	4	7
421	13	7	10	3	6	8	5	6	10	4	8	10	8	5	7	17	4	12	7	5
275	4	4	6	10	10	4	6	7	5	5	4	6	5	10	4	14	8	8	5	3
168	6	5	8	7	11	6	7	8	9	7	9	9	9	9	11	5	10	11	8	9
360	9	13	11	11	5	11	10	9	6	8	10	5	10	12	6	21	12	13	12	15
324	5	11	9	8	8	9	8	10	8	12	6	11	7	11	12	15	6	9	9	10
73	10	8	5	5	7	10	15	4	12	15	11	7	12	6	9	19	17	5	6	6
454	8	12	7	12	9	7	4	12	7	9	12	14	6	13	10	18	11	10	10	11
306	12	9	12	13	12	12	11	11	13	16	7	8	14	8	8	10	16	7	11	8
166	14	16	14	15	18	16	19	16	15	11	20	15	17	17	18	20	13	16	16	12
30	11	14	13	16	13	13	13	13	11	17	13	13	15	7	14	13	5	6	13	13
228	18	19	17	9	14	18	12	14	14	10	18	17	24	19	16	12	15	19	24	22
63	20	10	19	18	15	14	20	17	20	14	19	23	11	14	13	16	21	15	14	14
89	15	20	16	14	16	15	23	15	19	18	15	19	16	18	17	23	20	17	15	18
287	21	15	15	20	21	21	24	18	18	19	23	18	23	15	15	6	19	20	23	23
385	17	25	24	24	24	24	16	25	25	22	16	21	25	21	21	24	24	18	25	16
126	16	18	21	23	23	23	21	23	23	20	22	22	18	23	23	8	18	25	17	21
257	23	24	22	19	20	17	17	19	17	25	17	20	22	16	19	7	22	14	18	17
342	22	22	23	21	19	19	14	20	16	21	21	16	21	25	20	22	14	23	21	24
212	19	17	20	17	17	20	18	22	21	13	14	12	20	22	24	11	9	21	19	19
448	24	23	18	25	25	22	22	24	22	23	24	24	13	24	22	9	25	22	20	20
8	25	21	25	22	22	25	25	21	24	24	25	25	19	20	25	25	23	24	22	25

TABLE IV

NUMBER OF "BETTER" JUDGMENTS OF 20 JUDGES WITH A SET OF 25 SAMPLES

The Table Should Be Read, "One of the Judges Considered Sample 312 Better than Sample 289; None of the Judges Considered Sample 289 Better than Sample 303," etc.

Samples	289	303	195	421	275	168	360	324	73	454	306	166	30	228	63	89	287	385	126	257	342	212	448	8
312	1	0	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....
289	....	0	1	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....
303	....	....	3	1	2	1	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....
195	....	....	....	5	6	2	0	2	3	1	....	....	....	....	....	....	....	....	....	....	....	....	....	....
421	....	....	....	....	13	8	5	7	5	6	5	0	3	1	....	....	....	....	....	....	....	....	....	....
275	....	....	....	....	....	3	2	3	6	2	3	0	4	2	0	....	....	....	....	....	....	....	....	....
168	....	....	....	....	....	....	4	7	10	7	6	0	3	0	0	....	....	....	....	....	....	....	....	....
360	....	....	....	....	....	....	....	14	10	11	7	2	5	2	3	....	....	....	....	....	....	....	....	....
324	....	....	....	....	....	....	....	....	11	7	7	1	4	2	1	....	....	....	....	....	....	....	....	....
73	....	....	....	....	....	....	....	....	....	8	5	2	4	4	3	0	1	....	....	....	....	....	....	....
454	....	....	....	....	....	....	....	....	....	....	9	0	5	2	2	0	1	0	....	....	....	....	....	....
396	....	....	....	....	....	....	....	....	....	....	....	2	5	3	2	0	1	0	1	....	....	....	....	....
166	....	....	....	....	....	....	....	....	....	....	....	....	17	9	10	7	4	2	1	5	....	....	....	....
30	....	....	....	....	....	....	....	....	....	....	....	....	....	4	4	1	1	0	1	1	0	....	....	....
228	....	....	....	....	....	....	....	....	....	....	....	....	....	....	12	9	7	4	5	8	4	8	4	2
63	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	8	5	4	4	6	4	7	2	0
89	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	8	2	4	6	5	7	3	0
287	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	5	7	9	9	11	4	2
385	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	13	15	15	14	8	9
126	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	13	11	15	5	5
257	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	9	10	4	3
342	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	15	7	3
212	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	4	3
448	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	6
8	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....	....

From the data of Table IV the scale was derived by the method of right and wrong judgments.[1] If we compare sample 195 of this table with sample 421, it will be noticed that 5 of the judges considered sample 195 better than sample 421, and therefore the rest of the judges, 15 in number, considered sample 421 better than sample 195. The 5 judgments making sample 195 better than sample 421 represent the "wrong" judgments, and the 15 judgments making sample 421 better than sample 195 represent the "right" judgments. The data of Table IV are readily reducible to percentages of "better" judgments on inspection. For example, the comparison between sample 195 and sample 421 may read, "25 per cent of the judges considered sample 195 better than sample 421, and 75 per cent considered sample 421 better than sample 195."

The unit of difference between the samples of this scale is defined as that difference which 75 per cent of the judges are able to distinguish and 25 per cent of the judges fail to distinguish. If all of the judges consider one sample better or worse than another, it is evident that we have no means of knowing how much better or worse the one is than the other. Further, if half of the judges consider one sample better than another and half consider it worse, it is evident that the samples may be regarded as of equal value. When 75 per cent of the judges judge one composition better than another and 25 per cent regard it as worse, comparison of the surfaces of distribution of the judgments on the two compositions shows clearly that the median values of the two compositions are separated on a common scale by a distance equal to the probable error. A unit on the present scale is therefore 1 P.E.

A scale is more convenient as an instrument of measurement if the steps are of equal size. An inspection of Table IV shows that it is not possible to select from the samples therein represented all the samples of a scale so that each step will be exactly 1 P.E. Percentages of "worse" judgments other than 25 may be expressed in terms of probable error by reference to Table V.

If we begin the selection of a scale with sample 195 (Table IV), sample 421 will be the next member, one step higher, since 25 per cent of the judges considered sample 195 better than sample 421. In the same manner sample 360 was selected as the next higher member, since the difference between samples 421 and 360 was found to be 1 P.E. Sample 30 followed sample 360 in like fashion. Since no sample was found which differed from sample 30 by exactly 25 per cent worse judgments, sample 228 was selected because its percentage of worse judgments approximated the desired percentage as closely as that of any other sample. Sample 30 was considered better than sample 228 by 20 per cent of the judges. This difference in percentage expressed in terms of probable error, according to Table V, is 1.25 P.E. The remaining samples of the scale were selected by proceeding in the same manner. Table VI gives a list of these samples with the percentages of better judgments, the probable error differences, and the probable error distances from the lowest sample in the scale.

---

[1] G. S. Fullerton and J. M. Cattell, "On the Perception of Small Differences," Philosophical Series of the *Publications of the University of Pennsylvania*, 1892, No. 2, pp. 12 f.

TABLE V

TABLE FOR DETERMINING THE PROBABLE ERROR FROM THE PERCENTAGE OF WRONG JUDGMENTS

Percentage of Wrong Judgments	Probable Error	Percentage of Wrong Judgments	Probable Error	Percentage of Wrong Judgments	Probable Error
1...........	3.45	18...........	1.36	35...........	0.57
2...........	3.05	19...........	1.30	36...........	.53
3...........	2.79	20...........	1.25	37...........	.49
4...........	2.60	21...........	1.20	38...........	.45
5...........	2.44	22...........	1.14	39...........	.41
6...........	2.31	23...........	1.10	40...........	.38
7...........	2.19	24...........	1.05	41...........	.34
8...........	2.08	25...........	1.00	42...........	.30
9...........	1.99	26...........	0.95	43...........	.26
10...........	1.90	27...........	0.91	44...........	.22
11...........	1.82	28...........	0.86	45...........	.19
12...........	1.74	29...........	0.82	46...........	.15
13...........	1.67	30...........	0.78	47...........	.11
14...........	1.60	31...........	0.74	48...........	.07
15...........	1.54	32...........	0.69	49...........	.04
16...........	1.47	33...........	0.65	50...........	0.00
17...........	1.41	34...........	0.61	....	....

In addition to the percentages of better judgments from Table IV, there are presented in Table VI the percentages of better judgments on the same samples from Table II. We have in these two sets of figures a means of comparing the rating given these samples by the 20 students of education with that given by the like number of special teachers of English. It is interesting to note that not only do the samples preserve the same relative rank in both sets of judgments, but—what is more interesting—the intervals between samples are judged exactly the same by both in all but two cases, and in each of these cases the variation is slight. With so small a difference in the results from the two sets of judges, the writers seemed justified in adopting for the final scale the values from Table IV.

TABLE VI

VALUE INTERVALS BETWEEN SCALE SAMPLES

Samples	Percentage of Better Judgments, Table IV	P.E. Distance from Preceding Sample	Distance from Sample 195	Percentage of Better Judgments, Table II
195	....	....	0.0	....
421	75	1.0	1.0	60
360	75	1.0	2.0	75
30	75	1.0	3.0	75
228	80	1.25	4.25	80
126	75	1.0	6.25	75
448	75	1.0	7.25	85
8	70	0.78	8.03	70

Any scale to be of fullest service must be graduated with reference to an established zero point. From Table VI one can see that sample 8 is 5.03 steps above sample 30, but he cannot say how many times better sample 8 is than sample 30. In other words, the measurement is only relative. Another set of judgments was therefore necessary to determine the distance of the lowest sample of the scale from zero. A set of 7 compositions was accordingly submitted to 20 competent judges with instructions to arrange them in the order of merit. This list of 7 was made up of samples 580, 595, and 607 from the Hillegas list and samples 312, 287, 303, and 195 from our sixth-grade list. Sample 580, assigned zero value by Hillegas, was assumed here to have zero value. The other two Hillegas samples were selected because they were known to have values in the neighborhood of zero. The samples chosen from our sixth-grade list were those which, according to our figures, approximated most closely zero values. The results of these judgments appear in Tables VII and VIII.

TABLE VII

THE POSITIONS ASSIGNED TO 7 SAMPLES BY 20 JUDGES

Judges	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
670*	1	1	1	2	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1
580*	2	3	2	3	2	2	4	2	2	6	3	3	2	3	2	2	3	3	2	3
312	3	2	3	1	3	4	3	3	4	2	2	4	3	2	3	3	2	2	3	2
595*	5	7	4	4	5	3	1	4	5	7	5	2	7	4	4	4	4	5	4	5
289	7	4	7	5	6	6	7	6	3	5	4	7	6	5	6	5	4	7	7	4
303	4	5	5	7	4	5	6	7	7	4	7	5	5	7	5	6	6	4	6	7
195	6	6	6	6	7	7	5	5	6	3	6	6	4	6	7	7	5	6	5	6

*Samples selected from the Hillegas list.

In Table VIII it will be noticed that 45 per cent of the judges considered sample 580 better than sample 312. According to Table VI this percentage of worse judgments implies a difference between the two compositions of 0.19 of a step. Assuming the value of sample 580 to be zero as above stated, the absolute value of sample 312 is therefore 0.19 of a step above zero. Similarly the difference between samples 312 and 595 is found to be 1.54 P.E. The percentages of better judgments, the P.E. distances between the samples, and the P.E. distances above sample 580 appear in Table IX. The Hillegas samples do not appear in the final scale.

TABLE VIII

NUMBER OF BETTER JUDGMENTS FOR A SET OF 7 SAMPLES RANKED BY 20 JUDGES

Samples	580	312	595	289	303	195
670	0	1	1	......	......	......
580	......	9	2	1	1	1
312	......	......	3	1	0	0
595	......	......	......	6	5	3
289	......	......	......	......	11	9
303	......	......	......	......	......	11
195	......	......	......	......	......	......

TABLE IX

VALUE INTERVALS OF SCALE SAMPLES FROM AN ESTABLISHED ZERO

Samples	Percentage of Better Judgments	P.E. Distance from Preceding Sample	Distance from Sample 580	Final Value Assigned
580*	....	....	0. 0	0.0
312	55	0.19	0.19	0.2
595*	85	1.54	1.73	1.7
289*	70	0.78	2.51	2.5
195	55	0.19	2.70	2.7
421	75	1.00	3.70	3.7
360	75	1.00	4.70	4.7
30	75	1.00	5.70	5.7
228	80	1.25	6.95	7.0
126	75	1.00	7.95	8.0
448	75	1.00	8.95	9.0
8	70	0.78	9.73	9.7

* Not used in the final scale.

APPLICATION OF THE SCALE

The samples of the final scale, arranged in ascending order with the values affixed, were printed as they appear in the insert (facing p. 309) and used to score the samples of composition obtained from the 10 cities. Each sample of composition was scored by 5 competent individuals. The average of the 5 scores assigned to a sample was taken as the score of that sample. The average score of the samples from any grade was taken as the composition score of that grade. The results of the scoring appear in Table X. The different sixth grades, indicated by letters, appear in the order of rank according to their average scores. Rank 1 indicates the lowest rank and 10 the highest.

TABLE X

THE ABILITY IN ENGLISH COMPOSITION OF A SELECTED SIXTH GRADE IN EACH OF TEN MICHIGAN CITIES

City	Divisons of Grade	Score	Rank
A.	A	5.477	1
B.	A	5.582	2
C.	A and B	5.738	3
D.	A and B	5.941	4
E.	A	5.995	5
F.	A and B	6.077	6
G.	A	6.104	7
H.	A and B	6.196	8
I.	A and B	6.211	9
J.	A	6.446	10

An examination of the scores of Table X shows that the difference between the score of the lowest school on the list and that of the highest is 0.969, or practically 1 step on the scale. At first glance this difference may appear to be strikingly small, and one may be tempted to conclude that the composition abilities of the 10 schools are practically equal. The question might well be raised, however, How large a difference in ability is represented by one step on the scale? If one assume equal enrolments in the grades tested, and plot a surface of distribution of composition ability for School A and a similar one for School J, and if these surfaces of distribution be so placed that the median of one is 1 P.E.

distant from the median of the other, as the above difference would require, it will be observed that 75 per cent of the pupils of School J are above the median of School A. This means that for every two pupils of School A making a grade of 5.7 or better on the scale, there are three pupils in School J making the same grade or better. The meaning of a difference of one step on the scale might be expressed more clearly if we knew the amount of time or the amount of teaching effort that would be required to produce this degree of improvement in a sixth grade. In the opinion of the writers, both the time and the effort would probably be much larger than a cursory examination of the scores of Table X might lead one to suppose. It is evident that some three or four years of composition teaching has been necessary to advance the sixth grade of School A from a point somewhat above zero ability to an ability represented by the score 5.477. Again, the amount of time and effort necessary to advance a class one unit above 5.477 would certainly be greater than the amount of time and effort necessary to advance the same class a unit from any point below 5.477. Finally, judging from previous studies in measurement, the writers are of the opinion that the range of variation of median-grade abilities for the eight grades in composition would not be greater than the range of ability shown on this scale. In view of these considerations, it seems reasonable to suppose that there is in all likelihood as much as a year's composition training represented by the difference in the scores of Schools A and J. However, the amount of this difference could be determined by experiment with approximate accuracy.

To measure a grade by means of this scale, first administer the test according to the instructions given earlier in the paper. Then apply the scale to the results by comparing each composition carefully with the samples on the scale and recording for it as its score the numerical value of the scale sample resembling it most. The average of these scores will be the score of the grade. A comparison of this average with the scores of the schools presented in Table X should give a definite notion concerning the standing of the grade tested. The reader should be reminded, however, that the scores of these schools are probably somewhat higher than they would have been had the grades been randomly selected. They probably represent, therefore, a desirable standard of attainment or guide for sixth-grade classes.

In making comparisons two other points should be kept in mind: first, the tests were given near the middle of a semester, more accurately, the latter part of April; and secondly, 72 per cent of the pupils tested were in Grade VIA, that is, high-sixth grade. This was due, of course, to the fact that in several of the selected sixth grades there were no lower divisions.

It is hoped the scale may serve a wider usefulness than the measurement merely of sixth-grade composition ability. If one were desirous of sampling the composition product of a school system and were limited to the measurement of one grade, no grade, probably, would be more satisfactory for this purpose than the sixth. The sixth-grade results in this case might be regarded just as Stone regarded the results from his sixth-grade arithmetic tests—as symptomatic for the school system.

Our debt to Thorndike and Hillegas, which will be apparent throughout our paper, we gladly acknowledge.